

# 分割表項目数量化の荻野法に寄せて

水谷 静夫

## 0. 本稿の意図

本誌120集に載った荻野綱男の方法〔文献1〕は、属性的項目の分布を示す分割表における《最適尺度法》の一解法と言へる。荻野が出したやうな例に限らず、この方法が使へる問題は多い。原理的には、 $a \times b$ 分割表（即ちデータが $a$ 行 $b$ 列の行列として表示されるもの）の形にまとめるのが適当な場合には、この方法が利用出来る；但し $a$ や $b$ が少なくとも3でなければ、使っても実際の価値は乏しい。

荻野法の計算自体は別にむづかしくはないが、逐次近似計算なので手間が掛かる。該論文左15頁に言ふ通り、計算機やプログラム電卓を使はなければやり切れまい。そこで本稿では、先づ、これを計算する Basic プログラムを提示する。第二に、荻野の例より相当に大きい或計量語彙論の問題に適用した結果を報告する。第三に、荻野法を使ふに当っての若干の注意を述べる。以上が本稿の意図である。理論的な事柄は、予定されてゐる岩坪秀一の論文に待つ。

## 1. 荻野の例と水谷のプログラム

荻野が120集に表1として掲げたデータに対し、筆者の書いたプログラム CRQ 1 を走らせた出力が、図1である。実用上の収束判定条件を、繰返し計算の前回との差がすべて  $5 \times 10^{-4}$  未満とした；即ち小数第三位まで有効数字が得られるやうにした。ループ7回で6回目と有効数字が全く同じになり、従って数量化計算はこゝで打切られた。

図1の6回目の値を読者が120号の結果と比べる労を惜しまなければ、筆者の方がすべて正確に1だけ小さい事に気づくであらう。その理由は後に述べる。又、図1の出力には、荻野がしなかった、結果の規準化と最大化された積率相関係数の計算とが、加へてある。念の為に言へば、規準化とは、変量 $x, y$  それぞれに、平均が0で分散が1となるやうに変換を行ふ事である。この方法による結果は相対的なもの故、かゝる変換を施しても何ら本質的な違ひは生ぜず、かつ同類の他の結果と比べるのに便利なので、規準化する事はこの種の扱ひでは常套的な手段である。

さてこの計算をしたプログラム CRQ 1 を図2に掲げる。これは、ACOS 6（東芝）の対話型時分割システムの下で働く Basic で書いてある。やゝ冗漫な所もあるが、分りよさる旨として書いたつもりである。プログラムの構成は、

行0001-0040 制御用定数などの設定  
行0100 データ読み込み

(20) 分割表項目数量化の荻野法に寄せて

図1 荻野の例のCRQ1出力

```

          フンカツニヒョウ コウモク スウリョウカ (カクキョウ オキノホウ)
## 「コクゴ」ガク」 120シユウ ノ クイゴ」 スウリョウカ ##
Xノ シュウヘント」スウ: 1077 1146 436 290
Yノ シュウヘント」スウ: 485 501 492 488 493 490
テ-タ ノ オキキヤ: 2949

Xノ ショキチ          .000 1.000 2.000 3.000
=== 1 カイメ ノ Y          .000 .000 2.536 1.882 2.842 4.372 5.000
---- オアシク X          .000 1.830 2.529 3.000
=== 2 カイメ ノ Y          .000 .000 2.723 2.536 3.443 4.666 5.000
---- オアシク X          .000 1.987 2.605 3.000
=== 3 カイメ ノ Y          .000 .000 2.752 2.637 3.534 4.709 5.000
---- オアシク X          .000 2.011 2.617 3.000
=== 4 カイメ ノ Y          .000 .000 2.757 2.652 3.547 4.716 5.000
---- オアシク X          .000 2.015 2.619 3.000
=== 5 カイメ ノ Y          .000 .000 2.757 2.654 3.549 4.717 5.000
---- オアシク X          .000 2.015 .619 3.000
=== 6 カイメ ノ Y          .000 .000 .757 2.655 3.550 4.717 5.000
---- オアシク X          .000 2.015 .619 3.000
          スウリョウカ ノ クイサン シヤカン: 063 SEC

*** クック ノ キシユンカ ***

XCIC:--
-1.271 .477 1.000 1.331
YCIC:--
-1.482 -1.078 -.066 .412 1.034 1.185
ソカン クイスイ: .736

          キシユンカ ノ クイサン シヤカン: .007 SEC
    
```

- 行1000-1310 データ要目の計算・提示 (入力の正しさの確認用)
- 行2000-2910 数量化計算とループ各回の結果の提示
- 行3000-3600 規準化計算・提示 (相関係数も)
- 行4000-4199 印刷の下請
- 行4300-4399 最小値・最大値決定の下請

配列名と主な変数名との用途は、

- F: 分割表の各ますの値を入れる配列;
- X, Y: 第g回目の計算結果を入れる配列; U, V: 作業用配列;
- N1, N2: 行・列の周辺度数を記す配列; N: データの大きさ;
- L, S: 最大値・最小値; W, C, D: 作業用;
- M1, M2, D1, D2: x, yの平均値・標準偏差の算出用。

ここで、図1の結果が何故1だけ荻野の示したものと違ふのかを述べよう。一言で述べれば計算時間を短縮する為である。荻野の論文の左15頁③の式に見る通り、(中間)数値の算出に当ってそれぞれ、比例させた値に1を足してゐるが、この足し算はたゞ値を1からaまで又はbまでに配る為にだけ役立ち、最終結果に対してはそれ以外に何の貢献もしない。もしこれを0からa-1又はb-1までに配るとすれば、全くいらなくなる足し算である。荻野法で得られる結果は相対的なもの故、この変更はこの方法にとって本質的な変更ではなく、殊に結果を規準化してしまへば全く同一の数値に達する。図1の例で言っ

でも、言はばこの無駄な計算に  $(6+4) \times 7 = 70$  回のステップを費す。後述の図 4 の計算ではそれが約一万六千回にも上る。幾ら計算機にさせるからとて、無駄は無駄である。従ってこの改修を施した。もし荻野式の値が欲しければ、最終値に 1 を加へればよい。

なほ荻野は「普通 5 ~ 7 サイクルで値が収束する。」と言ふが、これは恐らく彼が手掛けた寸法の小さい分割表の場合の事である。(この点は後で再び取り上げる。) 寸法がもっと大きい分割表データを処理するには、一々中間出力をする CRQ 1 は不適當であり、又データの与へ方も行 5001-5004 の形式に限っては、使ひ勝手が頗る宜しくない。そこで別に CRQ 2 が作つてあるが、紙幅を慮つてそのプログラムリストは割愛する。次の節に述べる例は、そちらで実行したものである。

## 2. 計量語彙論からの一例

水谷が文献 2 で扱つた問題に荻野法を使つてみよう。その問題とは次のやうなものである。割り切つて言つてしまへば飽かず別れた女を後日追ひ求める唄である『湯の町エレジー』と『上海帰りのリル』とではあるが、しかし、曲調の事を別にしても、この二つから受ける我々の感じは著しく違ふ。その違ひが用語面から客観的に捉へられないか。この二篇に周辺の十八篇を併せ、その用語状況を林知己夫の数量化理論 第 IV 類・第 III 類の方法によつて分析し、成功と言へる結果を導いたのが、文献 2 である。

荻野法は或意味で第 III 類の方法に似た簡略版である。(双対的だからどちらを  $x$  と見てもよいが) キー語を  $x$ 、作品を  $y$  として両者を同時的に数量化しようとする。数量化第 III 類ではキー語出現の有無のパターンを手掛りとして作品の仕訳を企てる。ここではキー語の使用度数の多寡は捨象されてゐる。荻野式によれば使用度数の情報まで仕訳に参与させる事が出来る。だからと言つてどちらの方法が圧倒的に優れてゐるとは断難し難い。これは数理ばかりで決定出来る性質の問題ではない。計量語彙論の方法論の問題として大変に興味深くはあるが、本稿ではその議論を保留し、この類ひの分析にも荻野法の使途がある事だけを例示するにとゞめる。

図 3 はこの例題のデータであるが、見易さを考へて、CRQ 2 で計算した図 4 に示す結果に基づいて排列してある。荻野法(に限らずこの類ひの分析法)は、要するに、素データから図 3 のやうな一層整つた形のデータ排列を得、更にそこに隠れた秩序を読み取る為の方法とも言へよう。図 3 の右側にはキー語に、下側には作品に、この方法で与へられた値(即ち図 4 の規準化された  $x_i, y_j$ )を書き添へておいたが、その下側の値の組に見る通り、想定した作品群仕訳に成功してゐる。

こゝに挙げた 20 篇だけを考へてみても、およそ歌謡曲におのづからなる線型關係は期待しにくく、それらは多側面的な存在と見られるであらう。にも拘らずキー語をうまく選ぶ事によつて、0.834 といふ相當に高い積率相関係数を有つやうな、しかも我々の直観的な綜合判定による分類を裏付ける、一次元的排列が出来た事は、注目し得る。荻野法によるこの結果は、文献 2 での分析結果と大局的に合致する。相異なる三通りの方法でほぼ同じ結果が導けたから、これら歌謡曲の用語法にかなり安定した秩序・構造の存在を考へる事は、強ち無理でない。

図2 プログラムCRQ1

```

0001 PRINT "      プランツ_ヒョウ コウモク スクリョウカ (カクチョウ オキ)ノ=ホウ)"
0002 REM
0003 T$="「コクコ」カク 120シユウノ クイコ" スクリョウカ"
0004 PRINT "## ";T$;" ##"
0005 PRINT
0010 DIM F(4,6),N1(4),N2(6), X(1,4),Y(6,1), U(4,1),V(1,6)
0020 LET A=4 ¥ B=6 ¥ E=.0005
0030 AO=A-1 ¥ BO=B-1
0040 G=1 ¥ N=ZO=0
0100 MAT READ X,F
0110 MAT Y=ZER
1000 PRINT "X ノ ショウケン ト スウ:";
1010 FOR I=1 TO A
1020 W=0
1030 FOR J=1 TO B
1040 W=F(I,J)+W
1050 NEXT J
1060 M1(I)=W ¥ PRINT W;
1070 N=W+N
1080 NEXT I
1090 PRINT
1100 PRINT "Y ノ ショウケン ト スウ:";
1110 FOR J=1 TO B
1120 W=0
1130 FOR I=1 TO A
1140 W=F(I,J)+W
1150 NEXT I
1160 N2(J)=W ¥ PRINT W;
1170 NEXT J
1180 PRINT
1190 PRINT "テ-7 ノ オキマ:";N
1200 PRINT
1300 PRINT "X ノ ショキマ";TAB(16);
1310 GOSUB 4110
1990 Z=TIM(A)
2000 MAT V=X*F
2010 L=0 ¥ S=99999999
2020 FOR J=1 TO B
2030 W=V(1,J)/N2(J)
2040 GOSUB 4300
2050 V(1,J)=W
2060 NEXT J
2100 K=BO/(L-S)
2110 D=1
2120 FOR J=1 TO B
2130 W=(V(1,J)-S)*K
2140 IF ABS(Y(J,1)-W)<E THEN 2160
2150 D=D
2160 Y(J,1)=W
2170 NEXT J
2240 ZO=TIM(B)-Z+ZO
2250 GOSUB 4000
2900 IF C*D=1 THEN 2980
2910 G=G+1 ¥ GOTO 1990
2980 PRINT TAB(15);"スクリョウカ ノ クイケン シ"カン:";INT(1000*ZO)/1000;"SEC"
2990 PRINT
3000 PRINT "      *** ケツカ ノ キシ ユンカ ***"
3010 M1=M2=D1=D2=0
3020 Z=TIM(E)
3100 FOR I=1 TO A
3110 W=N1(I) ¥ C=X(1,I)
3120 W=W*C ¥ M1=M1+W
3130 W=W*C ¥ D1=D1+W
3140 NEXT I
3150 M1=M1/N
3160 D1=SQR(D1/N-M1*M1)
2490 Z=TIM(C)
2500 MAT U=F*Y
2510 L=0 ¥ S=99999999
2520 FOR I=1 TO A
2530 W=U(I,1)/N1(I)
2540 GOSUB 4300
2550 U(I,1)=W
2560 NEXT I
2600 K=AO/(L-S)
2610 C=1
2620 FOR I=1 TO A
2630 W=(U(I,1)-S)*K
2640 IF ABS(X(1,I)-W)<E THEN 2660
2650 C=C
2660 X(1,I)=W
2670 NEXT I
2740 ZO=TIM(D)-Z+ZO
2750 GOSUB 4100
3200 FOR J=1 TO B
3210 W=N2(J) ¥ D=Y(J,1)
3220 W=W*D ¥ M2=M2+W
3230 W=W*D ¥ D2=D2+W
3240 NEXT J
3250 M2=M2/N
3260 D2=SQR(D2/N-M2*M2)

```

← 標題を与へる

] 分割表の大きさ  
に合せて指定

## 図2 続き

```

3300 W=0
3310 FOR I=1 TO A
3320 K=X(1,I)
3330 FOR J=1 TO B
3340 W=K*Y(J,1)*F(I,J)+W
3350 NEXT J
3360 NEXT I
3370 W=(W/N-M1*M2)/D1/D2
3380 Z=TIM(F)-Z
3390 PRINT

3500 PRINT "XCII:--"
3510 FOR I=1 TO A0
3520 PRINT USING 3777, (X(1,I)-M1)/D1;
3530 NEXT I
3540 PRINT USING 3777, (X(1,A)-M1)/D1
      3550 PRINT "YCJJ:--"
      3560 FOR J=1 TO B0
      3570 PRINT USING 3777, (Y(J,1)-M2)/D2;
      3580 NEXT J
      3590 PRINT USING 3777, (Y(B,1)-M2)/D2

3600 PRINT "ツカノ ケヅ:";INT(1000*W)/1000
3610 PRINT
3620 PRINT TAB(15);"キ" ユカ ノ ケヅン シ カン:";INT(1000*Z)/1000;"SEC"
3630 STOP

      3777: ###.###

4000 PRINT "===";G;"カノ Y";TAB(24);
4010 FOR J=1 TO B0
4020 PRINT USING 3777, Y(J,1);
4030 NEXT J
4040 PRINT USING 3777, Y(B,1)
4099 RETURN

      4100 PRINT " ---- けし" X ";
      4110 FOR I=1 TO A0
      4120 PRINT USING 3777, X(1,I);
      4130 NEXT I
      4140 PRINT USING 3777, X(1,A)
      4199 RETURN

      4300 IF W>S THEN 4320
      4310 S=W
      4320 IF W<L THEN 4399
      4330 L=W
      4399 RETURN

5000 DATA 0,1,2,3
5001 DATA 438,371,164, 90, 10, 4
5002 DATA 39,110,280,279,243,195
5003 DATA 7, 17, 39, 88,142,143
5004 DATA 1, 3, 9, 31, 98,148
9999 END

```

データはその都度与へる

序でに、話がやゝ脇道に入るが、**数量化第 III 類**との関係について一言しておく。荻野法のデータ行列を作り替へて、 $f_{ij} \geq 1$  の所は改めて 1、 $f_{ij} = 0$  の所はそのまま 0 にすれば、かくて得られるデータ行列はまさに第 III 類のそれである。かやうに作り替へたデータに荻野法を適用すればどうなるかを実験的に確かめたところ、次のやうになった。(1) 荻野法で得られる相関係数は、第 III 類の方法での第一軸値に関する相関係数と等しい。(2)  $x_i$  の値は、時に符号が逆になる場合がある事を別にすれば、即ち絶対値としては、第一軸値と一致する。(3)  $y_j$  の値は、両方法の算出法の違ひのせいか、一致しない。この(3)に対する数理的究明をまだしてゐないが、荻野法は数量化第 III 類と縁が深い方法であると言へる。

図3 キー語の分布 附 荻野法による規準化した  $x_i$  及び  $y_j$

昭和	二六	二一	二七	二〇	二七	二四	二七	一四	一一	一一	一五	一四	一三	一三	二六	二六	二四	二二	二二	$z_i$	
	上海帰りのリル	上海リル	霧の透のリル	上海リル	私がリルよ	夜霧に消えたチャコ	ロンドン街角で	上海の花売娘	東京の花売娘	南京の花売娘	広東の花売娘	上海夜曲	上海ブルース	霧の四馬路	上海の街角で	湯の町物語	別府エレジー	湯の町夜曲	湯の町みかみうなぎ		
リル	2	1	18	4	10															-1.479	
ヒトリポッチ	1		1																	-1.421	
キ・クル	2	3	1						1											-1.282	
ドコ	1		2		1				1											-1.068	
オモイデ	1			2	1		1													-1.034	
ヒ				2	1	1														-0.977	
アカイ		2	1			1	2	1		1										-0.444	
キリ	1		3		1	1	2	1			1		1	1						-0.383	
ヒトミ	1	2		1			1	1	1	1		1	1							-0.333	
カナシイ			1		3	2							1		1	1	1			.022	
ユメ	1			2		1						1	1	1	1				1	.216	
ハナウリ							3	6	3		1									.306	
ハナカゴ							1	1	1	1										.321	
ムスメ							3	6	3	3	1									.327	
メシス							6		6		3									.339	
スレル							1			1	1									.377	
ハナ		1	1		1	1			6	12	6	2	1	1	1	1	1	1		.380	
カゲ							1	1	1			1								.381	
スマロ	1											1	1	2	1					.440	
ワカレル	1			1								2		1	3	2	1			.734	
チリル				1						2		1		1	1	1		1		.763	
ツキ					1		1	1		1		1	3	1	1	1	1	1	1	.886	
ムセビ・ア			1												1		1	1	1	.894	
コヨイ			1										1	1	1	1	1	1	1	.999	
ギミ												1	1	1	1	1	3	3	3	1.597	
ギタア										1							1	1	2	1.697	
タビ																1	1	1	1	1.792	
ユ															1	1	1	1	2	1.816	
(ラ)レル																1	1	1	1	1.869	
カ															1		1	1	1	1.922	
$z_j$	-1.410	-1.262	-1.229	-1.072	-0.918	-0.199	.157	.236	.330	.347	.441	.549	.609	.779	.896	1.009	1.418	1.528	1.666	1.675	$\rho = .834$

図4 キー語使用度数に基づく歌謡曲仕訳 (CRQ2の出力)

```

          フリガナコード コード名      スワリヨウカ (カクテヨウ オキ) ノ=ボウ)
### カコウキョウ:  ハナウリムスズ・リル・シヤンハイ・ユルムスチ ###
                                     ワイラン ツキ・ヒ・ネ 05/30/80 ノ 17.334シ
XCQJ: Δスズ  ハナウリ  ハナカコ  シ・ス  スレ=ル  カク
      リル  ヒ(DAY)  ヒトリホトチ  キ・クル  オモテ"  ト"コ
      1  タヒ"  カ(OODOR) (5)レ=ル  キ"ダア  Δセヒ"・フ
      アカ=イ  カガシ=イ  ナス  コヨイ  ツキ(MOON)  ハア
      ヌキ  ナリ(FOG)  ヒトヒ  スズヒ  カガレ=ル  ナリ・ル
YQJQ: 1-5 ハナウリムスズレノ 6-10 リルレノ 11 シ"ユルレノ
      12-15 ショウワ"10キ"ル"ダ"イ シヤンハイレノ 16-20 ユルムスチレノ

Xノシヨクノト"スウ:  16  13  4  15  3  5  4  61  4  10  3
      7  5  5  6  5  7  3  5  4  4  8  10
      7  14  35  13  12  10  6  11  3  16  31  22
Yノシヨクノト"スウ:  18  17  22  24  15  11  11  6  7  11
      30  7  10  11  10  9  15  11  6  7  11
テ"アノ"ノオキヤ:  310
Xノシヨクチ:
( 1) 1.000 ( 2) 2.000 ( 3) .000 ( 4) 6.000 ( 5) 3.000 ( 6) 4.000
( 7) 10.000 ( 8) 9.000 ( 9) 12.000 (10) 7.000 (11) 5.000 (12) 11.000
(13) 21.000 (14) 21.000 (15) 24.000 (16) 23.000 (17) 19.000 (18) 17.000
(19) 8.000 (20) 16.000 (21) 20.000 (22) 18.000 (23) 15.000 (24) 14.000
(25) 13.000 (26) 10.000 (27) 10.000 (28) 10.000 (29) 10.000 (30) 10.000
=== 316 カキテ XノA:
( 1) 15.492 ( 2) 15.317 ( 3) 15.440 ( 4) 15.599 ( 5) 15.926 ( 6) 15.960
( 7) .000 ( 8) 4.308 ( 9) .502 (10) 1.691 (11) 3.817 (12) -3.529
(13) 28.063 (14) 27.693 (15) 29.000 (16) 28.528 (17) 27.053 (18) 20.359
(19) 8.884 (20) 12.379 (21) 26.376 (22) 21.229 (23) 20.289 (24) 15.950
(25) 14.540 (26) 9.404 (27) 9.833 (28) 16.463 (29) 18.989 (30) 19.235
----- オアシ"ク YノA:
( 1) 10.141 ( 2) 11.399 ( 3) 10.818 ( 4) 10.718 ( 5) 9.649 ( 6) 2.081
( 7) .940 ( 8) .000 ( 9) 3.033 (10) 1.115 (11) 7.457 (12) 12.054
(13) 12.434 (14) 14.202 (15) 13.481 (16) 19.000 (17) 18.096 (18) 14.898
(19) 18.944 (20) 17.418

          スワリヨウカノワイランシ"カン:  12.658 SEC

          *** ケツカノキ"シ"ユカ ***
XCQJ:--
( 1) .327 ( 2) .306 ( 3) .321 ( 4) .339 ( 5) .377 ( 6) .381
( 7) -1.479 ( 8) -1.977 ( 9) -1.421 (10) -1.282 (11) -1.034 (12) -1.068
(13) 1.792 (14) 1.749 (15) 1.901 (16) 1.846 (17) 1.674 (18) .894
(19) -.444 (20) .022 (21) 1.595 (22) .995 (23) .886 (24) .380
(25) .216 (26) -.383 (27) -.333 (28) .440 (29) .734 (30) .763
YQJQ:--
( 1) .237 ( 2) .441 ( 3) .347 ( 4) .330 ( 5) .157 ( 6) -1.072
( 7) -1.262 ( 8) -1.410 ( 9) -.918 (10) -1.229 (11) -.199 (12) .549
(13) .609 (14) .896 (15) .779 (16) 1.675 (17) 1.528 (18) 1.009
(19) 1.656 (20) 1.418
          ヲウカンノ"スウ: .834
          シ"シ"ユカノワイランシ"カン: .087 SEC
  
```

図4に結果を示した計算を始めるに当って、素データ行列を一見しただけでは各キー語にどんな初期値を与えるのが良ささうか見当がつけ難かったので、図4に見るやうな、かなりでたらめな値の組を与へてみた。これは最終値を近似するやうなものではなかったの(実験の為わざとさうしてみたのであるが)、実用上の収束に至るまでに三百回余りの繰返し計算を要した。文献1に、初期値がどうであれ「わずかにサイクル数が1~2回多くなるだけ」と言ふのは、荻野の勇み足である。因みに、図4の最終結果を知った上で、規準化した  $x_i$  の小数第二位までの値(従って負数もある)を初期値として与へて験したとこ

## (26) 分割表項目数量化の荻野法に寄せて

ろ、今度は6回の繰返しで実用上の収束を見、この計算時間は0.2秒(図4のは12.7秒程)であった。

### 3. 荻野法における初期値の問題等

今述べた通り、荻野法を使ふ場合には初期値の与へ方に一考を要する。文献1に明記していないが、解法の形式的性質からして、

すべての初期値を同じ値にしてはならない。

その極端な場合として、すべてを0にすればどんな事が起るかを考へてみるがよい。かういふ初期値の組が無効である事は、すぐ分るであらう。又、 $\{0, 1, 2, 1, 1, 3, 1, 1\}$ のやうな等しい初期値の多用は、それらに対応するデータ分布が似てゐない限り、計算能率を落す原因となり勝ちである。

分割表の寸法が小さい時には、等しい値が多くても(すべてではない限り)苦になる程の計算時間の開きが出ない。もともとごく短時間に処理し切れるからである。しかし前節で例に挙げた程度の規模になると、大きな開きが出る。最終結果(の少なくとも大小順)を近似する初期値であれば、当然、計算時間が短くて済む。しかしどんなものが近似値なのか、一般には、事前に分らない。だからこそ計算に訴へる必要が有る訳である。それにしても、理由もなく変な値の組を使ふのは避けた方がよい。特に、消費時間による打ち切り制を採つてゐる計算センタを利用する場合には、かうした注意が必要であらう。

本稿の計算はすべて筆者の大学の中型機で行つたが、十秒を越す計算といふのは筆者自身は他に余りしてゐない。この点からすると、荻野法は解き方の直観的な分りよさにも拘らず能率的な算法ではないやうに思はれる。算法(algorithm)の改良が期待される。

結果の解釈法は計算方式とは別の事であり切り離して論ずべき事柄ながら、一つだけ言及しておかう。図3の「リル」は、特定作品にだけ集中してしかも著しく多く使はれてゐる点で、特異なキーである。さういふ特異なキーを含むと、相関関係が大きい割にすつきりした結果が得られない事が有る。文献2の25キー語を本稿では五語ふやして30としたが、25キーによる荻野法の結果では相関係数が0.883と大きかったのに作品仕訳としては鋭い結果にならなかつた;それで五つのキー語を追加して特異キーの影響を弱めようとしたのであった。一般には、相関係数の大小が数量化の成功の度合の指標になる。しかしさうでない場合が有る事を心得ておくのも無駄ではない。

### 4. 附言

計量国語学の分野では特に、具体的問題としては異なつても形式的には同じ処理法が反覆して使はれるのが普通である。さういふものをプログラム化して蓄積して行けば、単体としては簡単なものであつても幾つかをつないで使ふ事により案外の仕事が出来、研究能率が上がる。荻野法もさうした単体の一つとしてプログラミングした。筆者は、この学年の演習題目に計量語彙論支援プログラム集の作成を取り上げたが、本稿のプログラムはその教材の一つである。この演習で、実用に耐へるプログラムをせめて十種は作らせたいと思つてゐる。ユーザ層の拡大により使用体験からのフィードバックでプログラムの質を高

め範囲を拡げる途も拓ける。因みに数量化理論第 III 類を語彙論に適用する西村<sup>ひろ</sup>怨彦作 LIKE 3 の移植により、それを使って卒業論文を書いた又はまとめつゝある学生が既に四名あり、文法論・表記論等を含め計算機利用の卒業論文は本年度予定のものまで併せて十五六篇になる。上記の演習題目はかうした動きに支へられてゐる。文科系だからプログラムが書けないといふ事は無い。況や、自動車の運転と同様、作れない者には使へないといふ事は全く無い。但し工科系の使ひ方を徒らにまねるのでなく、文科系には文科系の使ひ様が有る点を銘記すべきであらう。国語学界にさういふ氣運が現れてもよい頃ではないか。本稿でプログラムを公開したのは、その呼び水になる事を欲してである。

なほ先に一寸触れた荻野法と数量化第 III 類との関係の考察には、線型代数学の知見がある。旧世代人である筆者はこの種の教育を受けて来なかつた。現代日本語に関する卒業論文の数が、理・工学部で文学部を上廻る勢ひである現況を見るにつけ、大学の国語学教育のカリキュラムが在来のまゝでいいかどうか、大いに氣懸りである。蛇足御審赦。

### 文献

- [1] 荻野綱男 (1980) 敬語における丁寧さの数量化——札幌における敬語調査から(2)——。『国語学』[120] 左13-24.
- [2] 水谷静夫 (1980) 用語類似度による歌謡曲仕訳、『湯の町エレジー』『上海帰りのリル』及びその周辺。『計量国語学』12 [4] 145-161.

——東京女子大学教授——